

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-086497

(43)Date of publication of application : 30.03.2001

(51)Int.Cl.

H04N 7/24

G06P 3/16

H04M 11/06

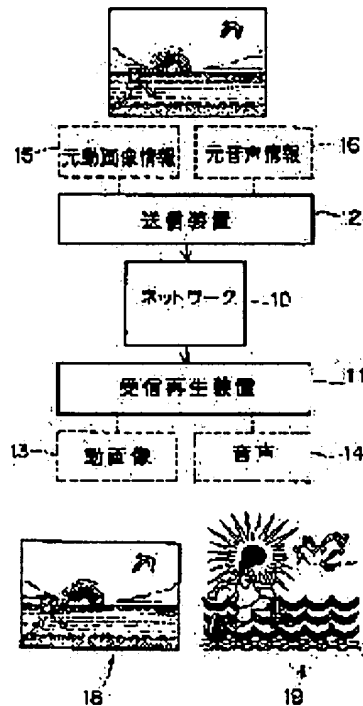
(21)Application number : 11-257647

(71)Applicant : MEGA CHIPS CORP

(22)Date of filing : 10.09.1999

(72)Inventor : NISHIMOTO MASAKAZU
KANEKO TOSHIKAZU

(54) INFORMATION COMMUNICATION SYSTEM



(57)Abstract:

PROBLEM TO BE SOLVED: To improve the qualities of reproduced images and reproduced sounds inside a receiving and reproducing device.

SOLUTION: The background or object and sound of original dynamic image information 15 and original sound information 16 are decomposed into respective elements, image recognition and sound recognition is performed for each of these element, and these elements are made into symbols, converted to transmission data and transmitted through a network 10 to a receiving and reproducing device 11. In the receiving and reproducing device 11, on the basis of the transmission data made into symbols, substitutive materials in a data base are read out in place of the respective elements of the original dynamic image information 15 and the original sound information 16, these substitutive materials are synthesized and a dynamic image 13 and a sound 14 are reproduced. At such a time, the motif of the substitutive materials is changed corresponding to the taste of a user and dynamic images 18 and 19 of various touches are selectively reproduced.

LEGAL STATUS

[Date of request for examination] 07.03.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3634687

[Date of registration] 07.01.2005

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-86497

(P2001-86497A)

(43) 公開日 平成13年 3月30日 (2001. 3. 30)

| (51) Int.Cl. ⁷ | 識別記号 | F I | テマコード (参考) |
|---------------------------|-------|---------------|-------------------|
| H 0 4 N 7/24 | | H 0 4 N 7/13 | Z 5 C 0 5 9 |
| G 0 6 F 3/16 | 3 4 0 | G 0 6 F 3/16 | 3 4 0 N 5 K 1 0 1 |
| H 0 4 M 11/06 | | H 0 4 M 11/06 | |

審査請求 未請求 請求項の数 6 O L (全 13 頁)

| | | | |
|-----------|--------------------------|----------|--|
| (21) 出願番号 | 特願平11-257647 | (71) 出願人 | 591128453 株式会社メガチップス 大阪市淀川区宮原4丁目1番6号 |
| (22) 出願日 | 平成11年9月10日 (1999. 9. 10) | (72) 発明者 | 西本 雅一 大阪市淀川区宮原4丁目5番36号 株式会 社メガチップス内 |
| | | (72) 発明者 | 金子 俊和 大阪市淀川区宮原4丁目5番36号 株式会 社メガチップス内 |
| | | (74) 代理人 | 100089233 弁理士 吉田 茂明 (外2名) |

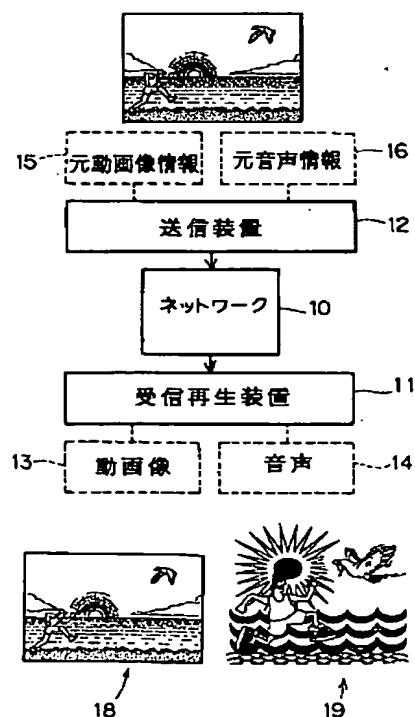
最終頁に続く

(54) 【発明の名称】 情報通信システム

(57) 【要約】

【課題】 受信再生装置内での再生画像及び再生音声の品質を向上する。

【解決手段】 元動画像情報15及び元音声情報16の背景やオブジェクト及び音声を個々の要素に分解し、これらの各要素毎の画像認識及び音声認識を行ってこれを記号化して送信データに変換し、ネットワーク10を通じて受信再生装置11に送信する。受信再生装置11では、記号化された送信データに基づいて、元動画像情報15及び元音声情報16の各要素に代えてデータベース内の代替素材を読み出し、この代替素材を合成して動画像13及び音声14を再生する。このとき、代替素材のモチーフを、ユーザの好みにより変更し、様々な画風の動画像18、19を選択して再生する。



(2)

【特許請求の範囲】

【請求項1】 送信装置により所望のコンテンツを所定のネットワークを通じて受信再生装置に送信または所定の記録媒体を介して運搬する情報通信システムであって、

前記送信装置は、

前記コンテンツに関する元動画画像情報及び元音声情報の少なくとも一方を個別の要素に分解する分解手段と、
前記分解手段で分解された個別の要素を認識する認識手段と、

前記認識手段での認識結果に基づいて前記各要素を記号化して送信データを生成する送信データ生成部と、
前記送信データ生成部で生成された前記送信データを所定のネットワークを通じて前記受信再生装置へ送信する送信側データ通信部、及び／または前記送信データ生成部で生成された前記送信データを所定の記録媒体に記録するデータ記録手段とを備え、

前記受信再生装置は、前記送信データを前記ネットワークを通じて受信または前記送信装置から運搬された前記記録媒体から読み出し、当該送信データに基づいて、所定のデータベース内に予め格納された代替素材を読み出し、これらの代替素材を合成して動画画像及び／または音声を再生するようにされたことを特徴とする情報通信システム。

【請求項2】 送信装置により所望のコンテンツを所定のネットワークを通じて受信再生装置に送信または所定の記録媒体を介して運搬する情報通信システムであって、

前記送信装置は、

前記コンテンツに関する元動画画像情報を単数または複数の背景画像及び単数または複数のオブジェクトに画像分解する画像分解部と、

前記画像分解部で画像分解された前記背景画像及び前記オブジェクトのそれぞれについて画像認識を行う画像認識部と、

前記コンテンツに関する元音声情報のなかから少なくとも主要な音声を抽出分解する音声分解部と、

前記音声分解部で音声分解された個々の音声データについて音声認識する音声認識部と、

前記画像認識部での認識結果及び前記音声認識部での音声認識結果に基づいて前記各背景画像、前記各オブジェクト及び前記各音声データを個別に記号化して送信データを生成する送信データ生成部と、

前記送信データ生成部で生成された前記送信データを所定のネットワークを通じて前記受信再生装置へ送信する送信側データ通信部、及び／または前記送信データ生成部で生成された前記送信データを所定の記録媒体に記録するデータ記録手段とを備え、

前記受信再生装置は、前記送信データを前記ネットワークを通じて受信または前記送信装置から運搬された前

2

記記録媒体から読み出し、当該送信データに基づいて、所定の画像音声素材データベース内に予め格納された代替素材を読み出し、これらの代替素材を合成して動画画像及び音声を再生するようにされたことを特徴とする情報通信システム。

【請求項3】 請求項2に記載の情報通信システムであって、

前記送信装置は、前記送信データ生成部で生成された送信データを編集する編集部をさらに備える情報通信システム。

【請求項4】 請求項2または請求項3に記載の情報通信システムであって、

前記受信再生装置の前記画像音声素材データベースは、前記送信データ中の前記背景画像、前記オブジェクト及び前記音声データの少なくとも一部について同一の前記送信データの情報に対応してそれぞれ複数種類の前記代替素材が用意され、

前記受信再生装置は、当該受信再生装置で動画画像及び音声を再生する際に、複数種類の前記代替素材の中から選択して前記背景画像、前記オブジェクト及び前記音声データの前記少なくとも一部のモチーフを変更することが可能とされたことを特徴とする情報通信システム。

【請求項5】 請求項3に記載の情報通信システムであって、

前記受信再生装置は、当該受信再生装置で動画画像及び音声を再生する際に、前記背景画像及び前記オブジェクトの少なくとも一部を変更または削除するための要求を所定の入力装置を通じて入力することが可能とされ、

前記送信装置の前記編集部は、前記受信再生装置の前記入力装置で入力された要求を認識し、当該要求を反映して前記送信データ生成部で前記送信データを生成するようにしたことを特徴とする情報通信システム。

【請求項6】 請求項2ないし請求項5のいずれかに記載の情報通信システムであって、

前記音声データの少なくとも一部は、前記背景画像または前記オブジェクトに対応付けられたリンク音声として認識されて、前記送信データ生成部で記号化される際にその対応付けについてのデータが含まれて前記送信データが生成され、

前記受信再生装置は、前記リンク音声の再生時に、前記背景画像または前記オブジェクトの再生に対応づけられるようにされたことを特徴とする情報通信システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、情報通信システムに関する。

【0002】

【従来の技術】近年の通信技術の発展に伴って、通信による様々な情報の提供サービスが試みられている。例えば一般公衆電話回線等を利用したインターネットを通じ

(3)

3

たデジタルコンテンツの提供もそのひとつである。

【0003】このようなデジタルコンテンツには、文字情報の他、音声情報や静止画及び動画を含む画像情報があり、例えば動画像である画像情報に合わせて音声情報が変化するような様々な組み合わせにより多様な情報提供を行うことが可能となっている。

【0004】

【発明が解決しようとする課題】ところで、インターネット等の一般公衆電話回線等のネットワークを通じて上記の多様な情報を提供する場合、ネットワークのデータ伝送速度の制限により、送受信される画像情報の画質及び音声情報の音質の向上は制約を受けてしまう。

【0005】通信によりリアルタイムで動画像を送受信する技術例としては、例えばISO/IEC JTC1/SC29/WG11によって国際標準化が進められている超低ビットレートに適したMPEG-4と呼ばれるデジタル動画像符号化方式が提案されている。このMPEG-4は、64 kbps以下の低ビットレートの動画像圧縮に適しており、ファイルサイズを抑え、オンラインでのストリーム再生や動画ファイルの転送を容易にできるものである。したがって、ISDN程度の回線速度があれば、これを受信した側でコマ落ちの少ない動画像として再生することが可能である。

【0006】しかしながら、このMPEG-4は、非可逆の圧縮方式であり、その圧縮率を極めて高くすることによって低ビットレート向きのデータに圧縮しているため、これを復号化して再生したときには相当な画質の劣化が生じる。具体的には、動画像の再生画質は、MPEG-1及びMPEG-2に比べると大幅に劣化してしまう。また、音声情報の音質劣化についても、これを非可逆の圧縮方式で送受信する限り、同様の問題が発生する。例えば、デジタルコンテンツにおいて人間が聴いて十分に高音質と感じる音声は、サンプリング周波数が44.1 kHzまたは音声の再生周波数帯域の上限が20 kHz前後で約6万5千階調(16ビット)以上であると言われている。しかしながら、現行の非可逆の圧縮方式でリアルタイムに音声情報を送信する技術では、このような高音質のリアルタイム再生を行うことはできない。

【0007】あるいは、これらの情報を、CD-ROMやICカード等の所定の記録メディア(記録媒体)を通じて送信することも考えられるが、この場合であっても、その記録メディアのデータ容量には限界があり、かかるデータ容量の制限内でしか情報を送信することができないという問題がある。

【0008】そこで、この発明の課題は、動画像情報や音声情報を含むデジタルコンテンツを送受信する場合に、受信先での再生時に画質及び音質を高い水準に保持できる情報通信システムを提供することにある。

【0009】

4

【課題を解決するための手段】上記課題を解決すべく、請求項1に記載の発明は、送信装置により所望のコンテンツを所定のネットワークを通じて受信再生装置に送信または所定の記録媒体を介して運搬する情報通信システムであって、前記送信装置は、前記コンテンツに関する元動画像情報及び元音声情報の少なくとも一方を個別の要素に分解する分解手段と、前記分解手段で分解された個別の要素を認識する認識手段と、前記認識手段での認識結果に基づいて前記各要素を記号化して送信データを生成する送信データ生成部と、前記送信データ生成部で生成された前記送信データを所定のネットワークを通じて前記受信再生装置へ送信する送信側データ通信部、及び/または前記送信データ生成部で生成された前記送信データを所定の記録媒体に記録するデータ記録手段とを備え、前記受信再生装置は、前記送信データを前記ネットワークを通じて受信または前記送信装置から運搬された前記記録媒体から読み出し、当該送信データに基づいて、所定のデータベース内に予め格納された代替素材を読み出し、これらの代替素材を合成して動画像及び/または音声を再生するようにされたものである。

【0010】請求項2に記載の発明は、送信装置により所望のコンテンツを所定のネットワークを通じて受信再生装置に送信または所定の記録媒体を介して運搬する情報通信システムであって、前記送信装置は、前記コンテンツに関する元動画像情報を単数または複数の背景画像及び単数または複数のオブジェクトに画像分解する画像分解部と、前記画像分解部で画像分解された前記背景画像及び前記オブジェクトのそれぞれについて画像認識を行う画像認識部と、前記コンテンツに関する元音声情報のなかから少なくとも主要な音声を抽出分解する音声分解部と、前記音声分解部で音声分解された個々の音声データについて音声認識する音声認識部と、前記画像認識部での認識結果及び前記音声認識部での音声認識結果に基づいて前記各背景画像、前記各オブジェクト及び前記各音声データを個別に記号化して送信データを生成する送信データ生成部と、前記送信データ生成部で生成された前記送信データを所定のネットワークを通じて前記受信再生装置へ送信する送信側データ通信部、及び/または前記送信データ生成部で生成された前記送信データを所定の記録媒体に記録するデータ記録手段とを備え、前記受信再生装置は、前記送信データを前記ネットワークを通じて受信または前記送信装置から運搬された前記記録媒体から読み出し、当該送信データに基づいて、所定の画像音声素材データベース内に予め格納された代替素材を読み出し、これらの代替素材を合成して動画像及び音声を再生するようにされたものである。

【0011】請求項3に記載の発明は、前記送信装置は、前記送信データ生成部で生成された送信データを編集する編集部をさらに備えるものである。

【0012】請求項4に記載の発明は、前記受信再生装

(4)

5

置の前記画像音声素材データベースは、前記送信データ中の前記背景画像、前記オブジェクト及び前記音声データの少なくとも一部について同一の前記送信データの情報に対応してそれぞれ複数種類の前記代替素材が用意され、前記受信再生装置は、当該受信再生装置で動画像及び音声を再生する際に、複数種類の前記代替素材の中から選択して前記背景画像、前記オブジェクト及び前記音声データの前記少なくとも一部のモチーフを変更することが可能とされたものである。

【0013】請求項5に記載の発明は、前記受信再生装置は、当該受信再生装置で動画像及び音声を再生する際に、前記背景画像及び前記オブジェクトの少なくとも一部を変更または削除するための要求を所定の入力装置を通じて入力することが可能とされ、前記送信装置の前記編集部は、前記受信再生装置の前記入力装置で入力された要求を認識し、当該要求を反映して前記送信データ生成部で前記送信データを生成するようにしたものである。

【0014】請求項6に記載の発明は、前記音声データの少なくとも一部は、前記背景画像または前記オブジェクトに対応付けられたリンク音声として認識されて、前記送信データ生成部で記号化される際にその対応付けについてのデータが含まれて前記送信データが生成され、前記受信再生装置は、前記リンク音声の再生時に、前記背景画像または前記オブジェクトの再生に対応づけられるようにされたものである。

【0015】尚、この明細書において、「コンテンツ」とは、画像及び／または音声等の種々の情報が制作者の意図をもって所定のレイアウトまたは所定のタイミングで編集された編集物をいい、また、「モチーフ」とは、各画像の形状、模様及び色彩や音声の音色、音程、和音、音声出現タイミング及びリズム等の種々の形態をいうものとする。

【0016】

【発明の実施の形態】＜原理＞デジタル情報の提供サービスのなかには、送信元に保有している動画像情報及び音声情報の元情報と意味内容が同等のものであれば、受信先で再生を行った際にその動画像及び音声のモチーフが異なっても差し支えない場合がある。むしろ、受信先のユーザの好みにより、元情報の動画像の画風等と異なる画風等に変更したいという場合さえある。このような場合は、元情報（動画像情報及び音声情報）のモチーフの完全同一性は重視されずにこの同一性の擬制が許容される一方、画質及び音質は高水準にしたいという要望がある。

【0017】例えば、図1のように、海を題材とするイメージ映像を送受信したい場合に、仮に元情報としての動画像（以下「元動画像」と称す）が実写映像であり、この元動画像内に含まれる空、海及び浜辺の風景を第一背景画像1とし、海の水平線の上に位置する夕日を第

6

二背景画像2とし、浜辺を左から右に向かって走る男性を第一オブジェクト3とし、空を右から左に向かって飛ぶカモメを第二オブジェクト4とした場合、これらの背景画像1、2及びオブジェクト3、4の位置関係及びその動きが受信側で高画質に再生されるならば、これらのモチーフが図2のように変更されても差し支えない場合がある。

【0018】また、音声情報についても同様である。例えば、図1に示した動画像情報の第一背景画像1（空、海及び浜辺の風景）に対応して波の音（第一背景音声）と風の音（第二背景音声）が聞こえ、また第一オブジェクト3（走る男性）の動きに沿って（リンクして）口笛（第一リンク音声）の音が左から右に移動し、さらに第二オブジェクト4（飛ぶカモメ）の動きに沿って（リンクして）カモメの鳴き声（第二リンク音声）の音が右から左に移動するような場合、これらの音声のそれぞれのモチーフの完全同一性は重視されずその擬制が許容される一方、音質については高音質の再生を行いたい場合がある。

【0019】このような場合において、この実施の形態の情報通信システムでは、動画像情報及び音声情報を含む元情報の各背景画像1、2中の移動物としての各オブジェクト3、4及びこれに付随する各リンク音声を所定の動き検出等の画像分解技術及び音声分解技術で抽出し、ここで抽出された各オブジェクト3、4及びリンク音声と、残りの各背景画像1、2及び各背景音声とをそれぞれ画像認識及び音声認識して記号化した後、この記号化された記号データのみをネットワークを通じて送受信するようにし、受信側では、受信した記号データに基づいて、情報格納装置内に予め用意しておいたモチーフの画像情報及び音声情報を呼び出し、これらを合成して再生（図2の例を参照）するようにしている。

【0020】以下、この情報通信システムを詳述する。

【0021】＜構成＞この情報通信システムは、図3の如く、デジタルコンテンツを所定のネットワーク10から受信して動画像情報13及び音声14として再生する受信再生装置11と、デジタルコンテンツとしての元動画像情報15及び元音声情報16を準備し且つネットワーク10を通じて受信再生装置11へ送信する送信装置12とを備える。

【0022】送信装置12は、図4の如く、元動画像情報15を背景画像及びオブジェクトに画像分解する画像分解部21と、画像分解部21で画像分解された背景画像及びオブジェクトのそれぞれについて画像認識を行う画像認識部22と、元音声情報16のなかから主要な音声を抽出して分解する音声分解部23と、この音声分解部23で音声分解された個々の音声データについて音声認識する音声認識部24と、画像認識部22での認識結果及び音声認識部24での音声認識結果に基づいて送信データを生成する送信データ生成部25と、送信データ

(5)

7

生成部25で生成された送信データを編集する編集部26と、送信データ生成部25で生成され編集部26で編集された送信データを含む種々の情報についてネットワーク10を通じて通信を行うデータ通信部27と、このデータ通信部27及びネットワーク10を通じて受信再生装置11から所望の要求があった際にその要求を認識する要求認識部28とを備える。

【0023】画像分解部21は、画像の周波数空間における高域フィルタを用いて画像中の各部分のエッジを抽出して領域抽出を行う領域抽出機能と、動画像中の各部分が時系列的に移動した際にMPEG等で使用される動き補償の方法を使用して画像中の各部分の動きを抽出する動き抽出（モーションディテクト）機能とを有し、動画像中で動いているオブジェクトを抽出するとともに、抽出されたオブジェクトを除いた部分を背景として、その背景をさらにいくつかの背景画像に分解するようになっている。尚、複数の合成された複合画像から個別のオブジェクト及び背景画像を抽出することが困難な場合もあり得るため、その場合には、元動画像情報15として複数のオブジェクト及び背景画像をいくつかのチャンネルに分けて用意しておき、これらの各チャンネルのオブジェクト及び背景画像を画像分解部21において要素認識するようにしてもよい。

【0024】画像認識部22は、図5の如く、画像分解部21で分解された各部分（背景画像及びオブジェクト）毎に、その特徴抽出31を行って所定の標準パターンに対してパターンマッチング32をし、その結果を送信データ生成部25に出力するようになっている。ここで、パターンマッチング32を行う場合は、音声認識部24で認識された音声の種類との間で関連づけを行うことが望ましいため、この関連づけ作業を状況参照33のブロックにて実行するようになっている。

【0025】音声分解部23は、図6の如く、元音声情報16のエンベロープを検出するエンベロープフォロー41と、エンベロープフォロー41からの出力信号を帯域毎に分割してそれぞれのピークを検出するピーク検出部42と、ピーク検出部42でのピーク検出結果からピッチ抽出を行うピッチ抽出ブロック43と、このピッチ抽出ブロック43で抽出された通過帯域を有するデジタルフィルタ44とを備え、特に、ピーク検出部42においてピーク検出される帯域として抽出したい音声の帯域（例えば人間の肉声等）を設定しておけば、波音、風の音、カモメの鳴き声及び口笛等の所望の音声を個別に抽出することが可能となる。尚、複数の合成された複合音声から個別の音声を抽出することが困難な場合もあり得るため、その場合には、元音声情報16として複数の音声データをいくつかのチャンネルに分けて用意しておき、これらの各チャンネルの音声を音声分解部23においてそれぞれ異なる音声として分解認識するようにしてもよい。

8

【0026】音声認識部24は、図7の如く、音声認識のための前処理を行うデータ前処理ブロック51と、この前処理されたデータに基づいて音声認識を行う認識部52とを備える。

【0027】データ前処理ブロック51は、音声分解部23で分解抽出された各音声データのエンベロープを検出するエンベロープフォロー55と、エンベロープフォロー55からの出力信号のピッチを平坦化して音量レベルを調整するレベル調整部56とを備え、認識部52での音声認識精度を向上するように前調整する。

【0028】認識部52は、ケプストラム変換や隠れマルコフ変換等の所定の変換処理を行う変換部57と、変換部57で変換されたデータについて音声の種類（波音や風の音等）を特定するとともに当該音声が人間の話す言語である場合に文字列に変換する音素片抽出部58とを備え、特に音素片抽出部58には音素片データベース（DB）59が接続され、この音素片DB59に基づいて音声の種類の特定または文字列の生成を行うようになっている。また、この認識部52は、認識した音声の時系列的な移動を認識する機能を有しており、この移動情報に基づいて、画像認識部22でパターンマッチング32を行った各背景画像またはオブジェクトとの間で関連づけを行い、これをリンク音声として認識するようになっている。例えば、認識部52で移動している音声があったときに、その移動している旨を画像認識部22の状況参照33のブロックに送信することで、動画像情報中のオブジェクトとの関連づけを行うことが可能であり、また音声が静止状態にある場合は、動画像情報中の背景画像との関連づけを行うことが可能となる。

【0029】尚、この認識部52は、データ前処理ブロック51から与えられた音声データがバックグラウンドミュージック（BGM）等の音楽である場合には、そのメロディ、ハーモニー及びリズムを検出し、その検出結果に基づいて当該音楽をデジタルデータとしてのMIDIデータに変換する機能をも有しており、これにより、記号化されたデータの他、MIDIデータをも併せて送信するようになっている。尚、音楽のなかからメロディ、ハーモニー及びリズムを検出することが困難な場合もあり得るため、その場合には、元音声情報16として各音色のMIDIデータをいくつかのチャンネルに分けて用意しておき、これらをそのまま用いるようにしてもよい。

【0030】送信データ生成部25は、画像認識部22での認識結果及び音声認識部24での音声認識結果に基づいて送信データを生成するものであり、所定のデータベース61内のデータを参照しながら各背景画像、各オブジェクト、各リンク音声及び各背景音声の記号化を行う。特に、動画像情報中の各オブジェクトまたは背景画像と音声情報中の個々の音声データとが関連づけられている場合に、これらをひとつの状況データとしてまとめ

(6)

9

10

て認識するようになっている。ここで、生成されたデータの例を次の表1に示す。

* 【0031】
* 【表1】

| 時刻 | 画像 | アクション | 音声 | アクション |
|-------|--------|--------|--------|---------|
| 00:00 | SUN000 | ACT000 | BGM010 | 0 |
| 00:05 | BOY008 | MOV104 | SNG408 | LINK-C3 |
| 00:12 | BRD425 | MOV203 | BRD443 | LINK-C4 |

【0032】ここで、表1中の時刻欄は各背景画像またはオブジェクトが元動画像情報において出現または新たな動作を開始した時刻を示している。また、表1中の画像欄は背景画像またはオブジェクトの種類を示す種類コードを示しており、例えば「SUN000」は夕日を、「BOY008」は走る少年を、「BRD425」は飛ぶカモメをそれぞれ意味する。さらに、表1中の左側のアクション欄は当該背景画像またはオブジェクトの動作を示す動作コードであり、「ACT000」は「SUN000（夕日）」が画面の中央部付近で赤みを帯びながら下降していく旨を、「MOV104」は「BOY008（走る少年）」が画面下部分で左から右に向けて移動する旨を、「MOV203」は「BRD425（飛ぶカモメ）」が画面上部を右から左に向けて移動する旨をそれぞれ示している。さらにまた、表1中の音声欄は音声の種類を示しており、「BGM010」はバックグラウンドミュージック（BGM）としての爽やかな音楽を、「SNG408」は口笛を、「BRD443」はカモメの鳴き声をそれぞれ意味している。そして、右側のアクション欄は、音声欄に記述された種類の音声の動きを示しており、「0」は静止状態を、「LINK-C3」は表1中の最上行から3行目の「BOY008（走る少年）」の動きにリンクするリンク音声である旨を、「LINK-C4」は最上行から4行目の「BRD425（飛ぶカモメ）」の動きにリンクするリンク音声である旨をそれぞれ意味している。

【0033】尚、この表1では、例えば「SUN000」などの種類コードを中心に画像や音声を表現しているが、実際に関連づけや受信再生装置11（後述参照）での代替画像の採択の際により詳細な情報が必要になる場合もある。このような場合には、種類コードに付随して、その種類コードで特定された要素の状態情報（例えば、「形状」、「形容」、「状況」等）を種類コードのサブセット（階層情報）として用意するなどしておけば、受信再生装置11での再生時に、より目的に合致した形で再生処理を行うことが可能となる。例えば、種類コードとして「SUN000」の太陽を特定した場合に、「形状」として「丸い」旨、「形容」として「赤い」旨、「状況」として「低い位置」にある旨をそれぞれコード指定した情報を生成する場合に、例えば、「SUN000」の情報を太陽から月に変更し、「形容」を「赤い」から「黄色い」に変更するなど、細やかな変更処理が容易になる。尚、これらのコードの内容は、原則

として受信再生装置11内に保存されているものと同様のものが使用される。ただし、受信再生装置11内に存在しない新しい情報を送信する必要がある場合には、その情報をコードと共に送信することも可能である。この場合、かかる新たな情報とコードとを、送信データの内部にまたは送信データの外部に関連づけて送信すればよい。

【0034】また、この送信データ生成部25は、音声認識部24からMIDIデータが与えられた場合には、このMIDIデータを上記の記号化されたデータに加えて送信データに含める機能を有している。

【0035】尚、この送信データ生成部25は、後述する編集部26での編集結果により、記号化して生成した送信データを変更する機能と、要求認識部28からの信号に基づいて記号化して生成した送信データを変更する機能とを有している。これにより、編集部26での編集結果や要求認識部28からの信号に基づいて、例えば上記の「BGM010」で表されるバックグラウンドミュージック（BGM）を省略するなどの個別の調整を行うことが可能となる。

【0036】編集部26は、送信データ生成部25で記号化されて生成された送信データに基づいて、所定のモニタ62に実際に動画像及び音声を再現し、これを見ながらキーボード及びマウス等の所定の入力部63を用いて送信データの変更及び削除等を行うようになっている。

【0037】データ通信部27は、送信データ生成部25で生成され編集部26で編集された送信データを含む種々の情報についてネットワーク10を通じて受信再生装置11へ送信する機能と、ネットワーク10を通じて受信再生装置11から与えられた要求信号を受信する機能とを有している。

【0038】要求認識部28は、データ通信部27及びネットワーク10を通じて受信再生装置11から要求信号が与えられたときにその要求を認識するもので、例えばバックグラウンドミュージック（BGM）を省略するなどの要求を送信データ生成部25に伝達するようになっている。

【0039】受信再生装置11は、上記の表1に示したように記号化された送信データを受信し、これに基づいて、コンテンツデータベース（DB）71、アクションルールデータベース（DB）72及び画像音声素材データベース（DB）73内に予め格納された代替素材を読

10

20

30

40

50

(7)

11

み出し、これらの代替素材を合成して動画像 1 3 及び音声 1 4 として再生するようになっている。

【0040】具体的に、この受信再生装置 1 1 は、ネットワーク 1 0 を通じて送信装置 1 2 との間の通信を行うデータ通信部 7 5 と、ネットワーク 1 0 及びデータ通信部 7 5 を通じて送信装置 1 2 から与えられた送信データ (表 1 参照) に基づいてコンテンツデータベース 7 1 内に格納されたコンテンツの種類 (画像であるか、音声であるか、文字であるか、あるいは音楽 (MIDI データ) であるか等) を特定するとともに、アクションルールデータベース 7 2 に格納された情報を参照して各オブジェクト等の動き (例えば左から右に移動している等) を置き換える置換処理部 7 6 と、置換処理部 7 6 で置き換えられた結果に基づいてデータ伸長すべきデータがあればデータ伸長を行うデータ伸長部 7 7 と、置換処理部 7 6 からの信号に基づいて各オブジェクト等の動きを算出するアクション算出部 7 8 と、アクション算出部 7 8 からの信号を画像音声素材データベース 7 3 内の代替素材に照合して各画像及び音声の素材を決定してレイヤーとして合成するデータ合成部 7 9 と、計時手段としての制御タイマー 8 0 と、制御タイマー 8 0 での計時に基づいてデータ合成部 7 9 からの出力画像及び出力音声の同期をとるためのシーケンサ 8 1 と、シーケンサ 8 1 からの出力に基づいて動画を再生する映像再生制御装置 8 2 と、映像再生制御装置 8 2 から出力された映像を最終的に合成する映像合成機 8 3 と、シーケンサ 8 1 からの出力に基づいて音声波形の合成を行うシンセサイザ 8 4 と、シンセサイザ 8 4 からの出力に基づいて音声合成を行う音声ミキサ 8 5 と、この受信再生装置 1 1 全体の制御を司る中央制御部 8 7 と、中央制御部 8 7 からの信号に基づいて送信装置 1 2 に対する要求を発行する要求部 8 8 とを備える。尚、シンセサイザ 8 4 は主として MIDI データの再生等を行うものであるが、例えば、音声情報として Wave データや MP 3 データ等を使用する場合は、これらのデータをシンセサイザ 8 4 を回避してシーケンサ 8 1 から音声ミキサ 8 5 へ直接送信するようになっている。

【0041】ここで、中央制御部 8 7 は、所定の入力装置 8 9 での入力に基づいて、受信再生装置 1 1 を使用するユーザーの希望事項を認識し、データ合成部 7 9 での画風 (例えば実写風であるか、漫画風であるか、あるいは、クラシック風であるか、現代ポップ風であるか等)、あるいは、バックグラウンドミュージック (BGM) が不要でない等の各種の設定を変更する機能を有しており、これらの情報は、要求部 8 8 を通じて送信装置 1 2 へ送信するようになっている。また、中央制御部 8 7 は、デジタルコンテンツを受信再生する用途 (例えば、業務用であるか、私用であるか等) についての情報や、使用している環境 (例えば、使用しているモデムや ISDN 等の通信速度) についての情報等の種々の情報

12

を要求部 8 8 を通じて送信装置 1 2 送信する機能を有している。

【0042】ここで、素材データベース 7 3 内の画像素材としては、静止画であってもよいが、各オブジェクトの表現で動きのあるものを表示したい場合を考慮して、複数の静止画を順次表示して動きある動画像を表示するアニメーション画像であってもよい。この場合、JPG やアニメーション GIF のような圧縮された画像素材を使用しても差し支えない。また、音声素材としても、MIDI データ、Wave データ及び MP 3 データのように、どのようなデータ形式のものを採用しても差し支えない。尚、例えば MP 3 データを使用する場合は、これらの圧縮されたデータを伸長するための伸長回路 (図示せず) を素材データベース 7 3 の内外に設置しておくことが望ましい。

【0043】また、素材データベース 7 3 内の代替素材 (画像素材及び音声素材) としては、それぞれ記号化された内容に対して 1 種類だけでなく、例えば、実写風、漫画風、クラシック風及び現代ポップ風のように、様々なユーザの好みの多様化を考慮して、それぞれの記号化された内容に対して多数のモチーフを用意しておく。

【0044】尚、上記した各構成要素は、例えば、専用の回路構成にてハードウェアとして構成されてもよく、あるいは、CPU を使用して所定のソフトウェアプログラムにしたがって動作する機能要素として実現しても良い。

【0045】<動作>上記構成の情報通信システムの動作を説明する。

【0046】まず、送信装置 1 2 側では、図 4 の如く、画像分解部 2 1 において、元動画像情報 1 5 を背景画像及びオブジェクトに画像分解する。この際、画像分解部 2 1 は、画像の周波数空間における高域フィルタを用いて画像中の各部分のエッジを抽出して領域抽出を行い、動き抽出 (モーションディテクト) 機能により動画像中の各部分の動きを抽出してオブジェクトを抽出するとともに、抽出されたオブジェクトを除いた部分を背景として、その背景をエッジ検出によりさらにいくつかの背景画像に分解する。尚、各オブジェクト及び背景画像が予め別々のチャンネルに分けて容易されている場合は、画像分解部 2 1 でそれぞれの要素認識を行うようにすればよい。

【0047】例えば、図 1 のように、海に沈みゆく夕日を題材とするイメージ映像であって、浜辺を左端から右に向かって少年が口笛を吹きながら走っており、空をカモメが右端から左に向かって飛んでいる元動画像情報 1 5 がある場合、画像分解部 2 1 は、この元動画像内に含まれる空、海及び浜辺の風景を第一背景画像 1 とし、海の水平線の上に位置する夕日を第二背景画像 2 とし、浜辺を左から右に向かって走る男性を第一オブジェクト 3 とし、空を右から左に向かって飛ぶカモメを第二オブ

(8)

13

ジェクト4とする。

【0048】次に、画像分解部21で画像分解された背景画像1、2及びオブジェクト3、4のそれぞれについて画像認識部22により画像認識を行う。この際、図5の如く、画像認識部22は、画像分解部21で分解された各部分（背景画像1、2及びオブジェクト3、4）毎に、その特徴抽出31を行って所定の標準パターンに対してパターンマッチング32をし、その結果を送信データ生成部25に出力するようになっている。

【0049】これにより、例えば図1の元動画像情報15については、第一背景画像1が空、海及び浜辺の風景であり、第二背景画像2が海の水平線の上に位置する夕日であり、第一オブジェクト3が浜辺を左から右に向かって走る少年であり、第二オブジェクト4が空を右から左に向かって飛ぶカモメである旨が認識される。

【0050】尚、この画像認識の際、後述の音声認識部24で認識された音声の種類との間で関連づけを状況参照33のブロックにて行うようにする。

【0051】これらの画像分解及び画像認識と併行して、音声分解部23により元音声情報16のなかから主要な音声を抽出して分解する。この際、音声分解部23は、図6の如く、エンベロープフォロワ41により元音声情報16のエンベロープを検出し、エンベロープフォロワ41からの出力信号をピーク検出部42で帯域毎に分割してそれぞれのピークを検出する。そして、ピーク検出部42でのピーク検出結果からピッチ抽出ブロック43がピッチ抽出を行い、ディジタルフィルタ44を通過させる。この際、ピーク検出部42においてピーク検出される帯域として抽出したい音声の帯域（例えば人間の肉声等）を設定しておく、波音、風の音、カモメの鳴き声及び口笛等の所望の音声を個別に抽出する。尚、各音声要素が予め別々のチャンネルに分けて容易されている場合は、音声分解部23でそれぞれの要素認識を行ってこれらを分解すればよい。

【0052】この音声分解部23で音声分解された個々の音声データは、音声認識部24で音声認識される。

【0053】この際、音声認識部24のデータ前処理ブロック51では、図7の如く、音声分解部23で分解抽出された各音声データのエンベロープをエンベロープフォロワ55で検出し、エンベロープフォロワ55からの出力信号のピッチをレベル調整部56で平坦化して音量レベルを調整する。その後、認識部52の変換部57においてケプストラム変換や隠れマルコフ変換等の所定の変換処理を行い、次に音素片抽出部58が、音素片DB59に基づいて、変換部57で変換されたデータについて音声の種類（波音や風の音等）を特定するとともに当該音声は人間の話す言語である場合にその言語の文字列の生成を行う。

【0054】また、この認識部52では、例えば口笛の音声左から右に向かって移動するような場合に、この

14

認識した音声の時系列的な移動を認識しておき、この移動情報に基づいて、画像認識部22でパターンマッチング32を行った各背景画像またはオブジェクトとの間で関連づけを行い、これをリンク音声として認識する。

【0055】尚、この認識部52は、データ前処理ブロック51から与えられた音声データがバックグラウンドミュージック（BGM）等の音楽である場合には、そのメロディ、ハーモニー及びリズムを検出し、その検出結果に基づいて当該音楽をディジタルデータとしてのMIDIデータに変換しておく。これにより、記号化されたデータの他、MIDIデータをも併せて送信することが可能となる。

【0056】そして、画像認識部22での認識結果及び音声認識部24での音声認識結果は、送信データ生成部25に送信され、これらに基づいて送信データが生成される。即ち、送信データ生成部25では、所定のデータベース61内のデータを参照しながら各背景画像、各オブジェクト、各リンク音声及び各背景音声の記号化を行い、上記した表1のようなデータ列としての送信データを生成する。特に、動画像情報中の各オブジェクトまたは背景画像と音声情報中の個々の音声データとが関連づけられている場合に、これらをひとつの状況データとしてまとめて認識するようにする。尚、表1中の各データの意味は上述したとおりである。

【0057】また、この送信データ生成部25は、音声認識部24からMIDIデータが与えられた場合には、このMIDIデータを上記の記号化されたデータに加えて送信データに含めるようにする。

【0058】ここで、作業者が編集部26で送信データの変更及び削除等を行った場合は、これに従って送信データが変更される。また、データ通信部27及びネットワーク10を通じて受信再生装置11から要求信号が与えられたときには、要求認識部28によりその要求を認識し、その要求を反映した送信データが生成される。これにより、例えばバックグラウンドミュージック（BGM）を省略するなどの要求に応じて、表1中の「BGM010」で表されるバックグラウンドミュージック（BGM）を省略するなどの個別の調整を行う。

【0059】このように生成された送信データは、データ通信部27からネットワーク10を通じて受信再生装置11へ送信される。

【0060】受信再生装置11では、上記の表1に示したように記号化された送信データをデータ通信部75で受信する。そして、置換処理部76は、データ通信部75で受信された送信データを読み取り、コンテンツデータベース71内に格納されたコンテンツの種類（画像であるか、音声であるか、文字であるか、あるいは音楽（MIDIデータ）であるか等）を特定するとともに、アクションルールデータベース72に格納された情報を参照して各オブジェクト等の動き（例えば左から右に移

(9)

15

動している等)を置き換える。

【0061】この置換処理部76で置き換えられた結果に基づいて、データ伸長部77は、データ伸長すべきデータがあればデータ伸長を行う。そして、アクション算出部78は、置換処理部76からの信号に基づいて各オブジェクト等の動きを算出する。データ合成部79は、アクション算出部78からの信号を画像音声素材データベース73内の代替素材に照合して各画像及び音声の素材を決定してレイヤーとして合成する。

【0062】そして、シーケンサ81により、制御タイマー80での計時に基づいてデータ合成部79からの出力画像及び出力音声の同期をとりながら、映像再生制御装置82で動画を再生し、映像合成機83で最終的な画像のレイヤー合成を行ってこれを所定のモニタ装置

(図示しない)に出力するとともに、MIDIデータの場合はシンセサイザ84で音声波形の合成を行って、音声ミキサ85で音声合成を行いこれを出力する。尚、音声情報としてWaveデータやMP3データ等を使用する場合は、これらのデータをシンセサイザ84を回避してシーケンサ81から音声ミキサ85へ直接送信して再生を行う。

【0063】ここで、受信再生装置11側のユーザが、その要望を所定の入力装置89で入力した場合、中央制御部87はそのユーザーの要望を認識し、動画像13及び音声14の合成に反映させる。

【0064】例えば、元動画像情報15と同一のモチーフの画像素材が画像音声素材データベース(DB)73内に予め格納されている場合であって、ユーザがそのままのモチーフを維持するように指定した場合は、図3中の符号18の如く、送信装置12で用意した元動画像情報15と殆ど同一の動画像13を受信再生装置11で再生することができる。このことは、音声14についても同様であり、ユーザが音声についてそのままのモチーフを維持するように指定した場合は、送信装置12で用意した元音声情報16と殆ど同一の音声14を受信再生装置11で再生することができる。

【0065】また、図1のように、元動画像情報15が実写風である場合に、ユーザが受信再生装置11において漫画風の再生を希望する場合は、このユーザが入力装置89を通じてその旨を中央制御部87に伝達する。この場合、データ合成部79は、中央制御部87からの指令に従って、画像音声素材データベース73内から読み出す各画像素材として、実写風のものを選択せずに漫画風のものを選択して読み出すようにする。これらの画像素材をデータ合成部79で合成した結果は図3中の符号19及び図2のようになる。この場合、図2の如く、空、海及び浜辺の風景である第一背景画像1、海の水平線の上に位置する夕日である第二背景画像2、浜辺を左から右に向かって走る少年である第一オブジェクト3、空を右から左に向かって飛ぶカモメである第二オブ

16

ジェクト4の全てにおいて、意味内容は同一であってもモチーフが元動画像情報15とは全く異なっており、これにより、ユーザの希望の画風に容易に変更することが可能となる。さらに、例えばユーザ自身の実写画像を画像音声素材データベース73内に用意しておけば、受信再生装置11での動画像13の再生時に、自分自身をモチーフにした動画像13を楽しむことができる。

【0066】ここで、表1の例などのように、一部のデータを他のデータにリンクさせる場合(例えば、音声データとしてのリンク音声をオブジェクトに対応させる場合)は、動画像と音声とが同期して動くなど、極めて自然な状態の対応づけが行われて再生されるため、高水準なデジタルコンテンツ提供サービスを行うことができる。

【0067】また音声14についても同様に、様々なモチーフを変更して楽しむことが可能である。例えば、送信装置12から与えられてきたMIDIデータをバックグラウンドミュージックとすることに代えて、画像音声素材データベース73内に予め格納されていた全く異なる音楽をバックグラウンドミュージックとしてもよい。これらは、ユーザが入力装置89に入力するだけで容易に変更できる。したがって、ユーザの好みに応じたデジタルコンテンツを容易に提供することが可能となる。

【0068】以上の動作において、ネットワーク10を通じて送信装置12から受信再生装置11に与えられるデータは、表1に示したような記号化された送信データが中心となっており、MPEG等の動画像そのものや、WAVEデータまたはMP3等の音声データそのものの通信を行う必要がないため、例えば、モデムを使用した一般公衆電話回線やISDN等を使用してインターネット上で通信を行うような場合でも、回線容量の限界により送信装置12から送信した情報の受信再生装置11側でのレスポンスの悪化を大幅に低減でき、極めて良好な通信レスポンスのデジタルコンテンツ提供サービスを実施することが可能となる。具体的には、元動画像情報15及び元音声情報16の情報量が数メガバイトであるような場合に、これをそのままネットワーク10を通じて送信すると通信時間が極めて長くなってしまい、リアルタイムの通信を行うことが不可能であるのに対し、この実施の形態では、この数メガバイトの情報量を数10バイトまで縮小して置き換えることができ、情報量が数10万分の1以上の圧縮効果を得ることができ、ネットワーク10に負荷をかけずに通信することが可能となる。

【0069】しかも、受信再生装置11の画像音声素材データベース73内の代替素材として高解像度の画像素材と高音質の音声素材を用意しておけば、上述のようにレスポンスの悪化を招かずに、極めて良好な画質で優れた音質のデジタルコンテンツを再生することが可能となる。

(10)

17

【0070】尚、デジタルコンテンツの編集は、図4のように送信装置12内の編集部26でも可能であるため、いわゆる常連のユーザが常に背景画像やバックグラウンドミュージックを省略するよう要求している場合などでは、予めそのユーザに対して背景画像やバックグラウンドミュージックを省略するよう編集部26で編集することも可能である。

【0071】また、背景画像やバックグラウンドミュージックがコンテンツにとって必須でない場合もあるため、各ユーザが低速のモデムを使用していることが要求認識部28で認識できた場合には、これに基づいて編集部26が自動的に背景画像やバックグラウンドミュージックを省略するように送信データ生成部25に指令するようにしてもよい。

【0072】さらに、送信装置12内の編集部26や受信再生装置11内での中央制御部87により、例えばクラシック調の画風を表現する際には、全体的に茶色っぽい深みのある色彩に統一するなどの色彩調整を行うようにしても良い。

【0073】尚、上記実施の形態では、受信再生装置11の画像音声素材データベース73において、全ての背景画像、全てのオブジェクト及び全ての音声データについて、異なるモチーフを予め用意し、これらの全てをユーザの希望に応じて変更できるようにしていたが、背景画像、オブジェクト及び音声データの一部についてのみユーザの希望に応じて変更できるようにしても差し支えない。この場合、背景画像、オブジェクト及び音声データの3つのカテゴリーのうちのいずれかひとつまたはふたつのカテゴリーのみを変更できるようにしてもよいし、あるいは、それぞれのカテゴリーについて一部の情報（背景画像、オブジェクト及び音声データ）のみを変更できるようにしてもよい。

【0074】また、上記実施の形態では、一般公衆電話回線等のネットワーク10を通じて送信データを送信装置12から受信再生装置11に与えていたが、CD-ROMやICカード等の記録メディア（記録媒体）を通じて送信データを受信再生装置11に与えるようにしても良い。この場合、送信装置12には、これらの記録媒体に送信データを記録するための記録装置（記録手段）を設置する必要があるが、また受信再生装置11側においても、これらの記録媒体から送信データを読み込むドライブ装置が必要となることは言うまでもない。このように、記録媒体を使用する場合であっても、元動画像情報15や元音声情報16等をそのままの状態を送信する場合に比べて、送信データその記録媒体のデータ容量の制限を受けずに済むという利点がある。尚、このように所定の記録メディアを使用する場合は、受信再生装置11側からユーザの要望を送信装置12に送信することが困難であるため、ネットワーク10を通じて通信と併用するすることが望ましい。

18

【0075】

【発明の効果】請求項1及び請求項2に記載の発明によれば、送信装置の画像分解部で元動画像情報を各背景画像及び各オブジェクト毎に分解するとともに、元音声情報の少なくとも主要な音声を音声分解部で抽出分解するなど、所定の分解手段によりコンテンツの要素分解を行い、それぞれ画像認識及び音声認識を行った後、その内容に応じて記号化して送信データを生成して受信再生装置に送信するようにし、受信再生装置側では、送信されてきた送信データに基づいて例えば所定の画像音声素材データベース内に予め格納された代替素材を読み出し、これらの代替素材を合成して動画像及び音声を再生するようにしているので、ネットワークまたは記録媒体を通じて送信装置から受信再生装置に与えられる記号化された送信データは、MPEG等の動画像そのものや、WAVEデータまたはMP3等の音声データそのものに比べて遙かに少量のデータになり、例えば、モデムを使用した一般公衆電話回線やISDN等を使用してインターネット上で通信を行うような場合でも、回線容量の限界により送信装置から送信した情報の受信再生装置側でのレスポンスの悪化を大幅に低減でき、極めて良好な通信レスポンスのコンテンツ提供サービスを実施することが可能となる。あるいは、記録媒体を通じて送信データを送信する場合であっても、その記録媒体のデータ容量の制限を受けずにすむという利点がある。

【0076】請求項3に記載の発明によれば、送信装置の編集部において、受信再生装置へ送信すべき送信データを編集することが可能となっているため、例えば受信再生装置側のユーザの希望に応じて、またはネットワークの通信速度等の様々な環境等に応じて、送信装置側でコンテンツの内容を変更することが可能となる。

【0077】請求項4に記載の発明によれば、ユーザの希望により、元動画像情報及び元音声情報と異なるモチーフの動画像及び音声を再生したり、一部の背景画像、オブジェクト及び音声を省略することが容易に可能となり、ユーザの希望に応じたコンテンツの提供を通信の負荷が増大することなく容易に実行することが可能となる。

【0078】請求項5に記載の発明によれば、ユーザの希望を、送信装置から送信される送信データに反映させることができるので、ネットワークに流れる送信データにユーザが希望しない無駄なデータが含まれるのを防止でき、効率のよい通信を実行することが可能となる。

【0079】請求項6に記載の発明によれば、音声データの少なくとも一部が、背景画像またはオブジェクトに対応付けられたリンク音声として認識されて、送信データ生成部で記号化される際にその対応付けについてのデータが含まれて送信データが生成され、受信再生装置は、リンク音声の再生時に、背景画像またはオブジェクトの再生に対応づけられるようにしているので、例え

(11)

19

ば、動画像中のオブジェクトとリンク音声とが同期して動くなど、極めて自然な状態の対応づけが行われて再生されるため、高水準なコンテンツ提供サービスを行うことができる。

【図面の簡単な説明】

【図 1】送信装置内に用意された元動画像情報の一例を示す図である。

【図 2】受信再生装置で再生する動画像の一例を示す図である。

【図 3】この発明の一の実施の形態に係る情報通信システムの全体構成を示すブロック図である。

【図 4】この発明の一の実施の形態に係る情報通信システムの送信装置を示すブロック図である。

【図 5】画像認識部を示すブロック図である。

【図 6】音声分解部を示すブロック図である。

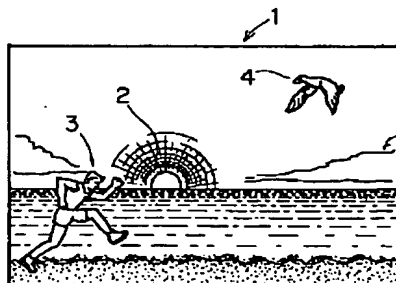
【図 7】音声認識部を示すブロック図である。

【図 8】この発明の一の実施の形態に係る情報通信システムの受信再生装置を示すブロック図である。

【符号の説明】

- 1, 2 背景画像
- 3, 4 オブジェクト
- 10 ネットワーク
- 11 受信再生装置
- 12 送信装置
- 13 動画像
- 14 音声
- 15 元動画像情報
- 16 元音声情報
- 21 画像分解部
- 22 画像認識部
- 23 音声分解部
- 24 音声認識部
- 25 送信データ生成部

【図 1】



20

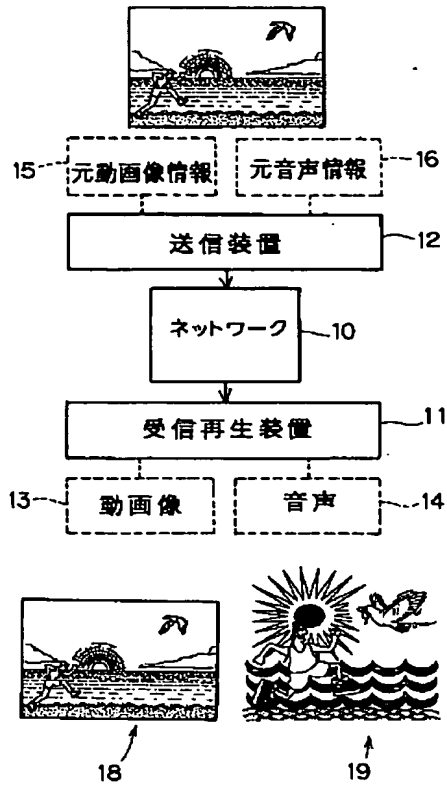
- 26 編集部
- 27 データ通信部
- 28 要求認識部
- 31 特徴抽出
- 32 パターンマッチング
- 33 状況参照
- 41 エンベロープフォロー
- 42 ピーク検出部
- 43 ピッチ抽出ブロック
- 44 デジタルフィルタ
- 51 データ前処理ブロック
- 52 認識部
- 55 エンベロープフォロー
- 56 レベル調整部
- 57 変換部
- 58 音素片抽出部
- 62 モニタ
- 63 入力部
- 73 画像音声素材データベース
- 75 データ通信部
- 76 置換処理部
- 77 データ伸長部
- 78 アクション算出部
- 79 データ合成部
- 80 制御タイマー
- 81 シーケンサ
- 82 映像再生制御装置
- 83 映像合成機
- 84 シンセサイザ
- 85 音声ミキサ
- 87 中央制御部
- 88 要求部
- 89 入力装置

【図 2】

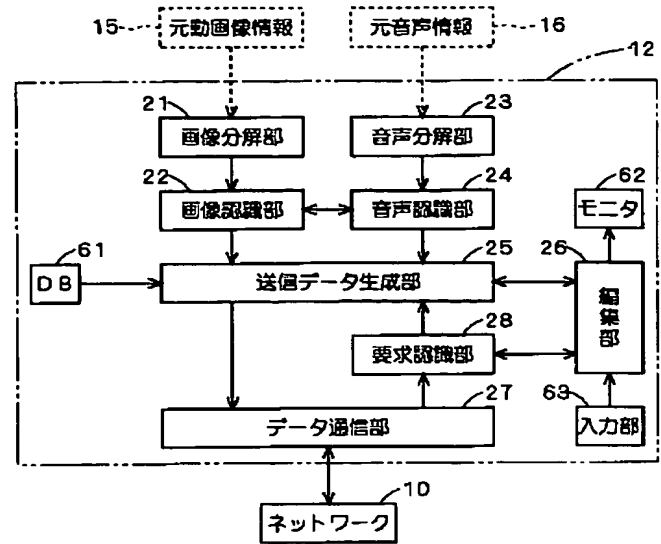


(12)

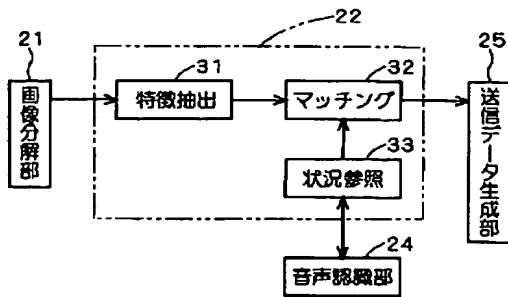
【図3】



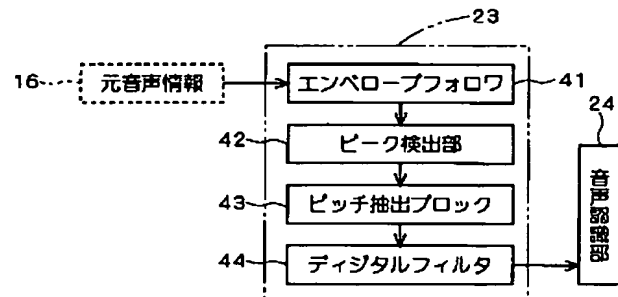
【図4】



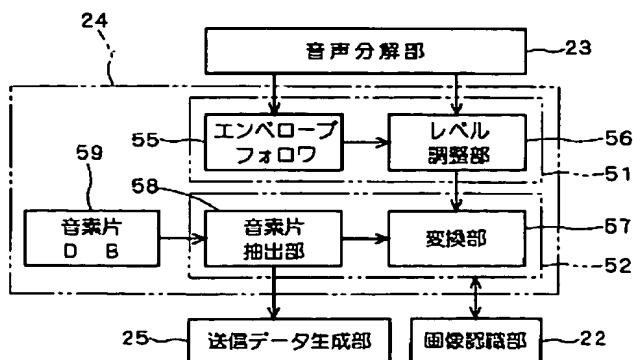
【図5】



【図6】

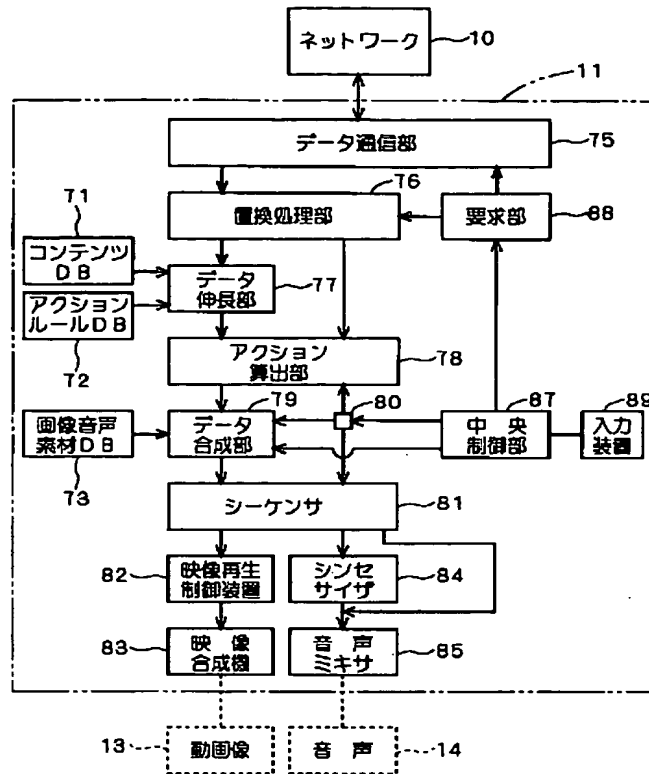


【図7】



(13)

【図8】



フロントページの続き

Fターム(参考) 5C059 KK01 KK11 KK37 LA01 LB11
 MB06 MB16 MB22 MB24 MB29
 PP10 PP26 PP27 PP29 RA01
 RD03 SS08 UA02 UA05 UA28
 UA31
 5K101 KK20 LL01 NN00 NN08 NN18
 NN21 TT06 UU19